Ethnicity – Reference Guide
Version 1.0, December 1999

Summary: Ethnicity

Definition

Ethnic background or association of a person.

Data Storage and Field Values

The Ethnicity data concept consists of a minimum of 1 data element, which serves as a 1-character variable indicator for ethnicity. The variable name noted below specifically applies to the Ethnicity of the subject of the report. Variable names for other uses of Ethnicity, such as Ethnicity of the subject's partner, are not discussed in this document but will be discussed in subsequent releases.

Variable Name:	HISPANIC
Type:	character
Length:	1
Reported to CDC:	Yes
Field Values:	Y – Yes (indicates *Hispanic or Latino)*
	N – No (indicates *Not Hispanic or Latino)*

An optional CIPHER extended ethnicity data element can be used to support collection of more detailed information, in standardized coded format, on ethnicity. NOTE: A table of extended ethnicity codes and categories titled "CDC/Census Ethnicity Classification", can be found on page 15 of this section. The optional CIPHER extended ethnicity race data element is described below:

> Variable Name: HISP_EXT
> Type: character
> Length: 6
> Reported to CDC: Yes
> Field Values: Refer to CDC/Census Ethnicity Classification in
>   this section

Missing Values

If the value of the Ethnicity data element is missing, or does not adhere to the CIPHER standard, the data element may be noted as blank to indicate a missing value. If the program requires the reason the value is missing, use a separate 1-character field to denote the rationale behind the missing data.

In addition, a separate 1-character field can be used for the optional extended Ethnicity data element if a program has chosen to make use of this element, and if the program requires collection of a reason that the value for an extended Ethnicity data element is missing.

The use of a Missing Value Reason data element must adhere to the CIPHER definition and rules associated with missing data as described in Appendix I - Missing Value Reason.

Processing Overview

Special requirements apply. Refer to the Implementation subsection on Data Processing: Validations and Edit Checks, below, for detailed information.

Discussion

The CIPHER Ethnicity data concept is influenced by the revised (1997) Federal Office of Management and Budget (OMB) Directive-15, and the Ethnicity standards supplied in the HISSB Standards and Liaison Committee's Common Data Elements Guide. The Ethnicity concept consists of one data element, which serves as a 1-character variable indicator for the OMB-defined minimum ethnicity categories of *Hispanic or Latino* and *Not Hispanic or Latino*. The data element, called HISPANIC, is formatted as a Yes/No

type variable in which valid values are 'Y' for Yes, and 'N' for No. A value of 'Y' in HISPANIC is an indicator of the ethnicity category *Hispanic or Latino*. A value of 'N' in HISPANIC is an indicator of the ethnicity category *Not Hispanic or Latino*.

Similarly, the collection of Ethnicity data is approached in a check-box type format in which one of the following ethnicity categories can be marked/checked off:

> *Hispanic or Latino*
> *Not Hispanic or Latino*

Selection of the category *Hispanic or Latino* results in the entry/storage of 'Y' in the HISPANIC data element. Selection of the category *Not Hispanic or Latino* results in the entry/storage of 'N' in the HISPANIC data element.

In addition, programs have the option to collect more detailed extended ethnicity information when the ethnicity category is noted as *Hispanic or Latino* (HISPANIC='Y'). These extended ethnicity categories are defined based on the revised (1997) Federal OMB Directive-15, and the Ethnicity standards supplied in the HISSB Standards and Liaison Committee's Common Data Elements Guide. As described in these existing standards, a unique 6-digit identifier (formatted as ####-#) is assigned/associated with each extended ethnicity category. The extended ethnicity data element (HISP_EXT) is activated only if the program has opted to collect/support these additional data and only if the associated Ethnicity data element (HISPANIC) is set to 'Y', indicating an ethnicity of *Hispanic or Latino*.

A table of extended ethnicity concept codes and categories supported within CIPHER (the CDC/Census Ethnicity Classification) is at the end of this section. The code set is compatible with but more detailed than the Race categories specified in the federal standards established by the U.S. Office of Management and Budget (OMB), and lends itself for use in electronic messages that comply with HL7 and other communications standards. The concepts included in the CDC/Census Classification table encompass the minimum OMB categories (described above) and a more detailed set of OMB-compatible categories provided by the U.S. Bureau of the Census. For each concept, a unique 6-digit identifier and a hierarchical code are provided. Synonyms are included as needed.

As described above, collection/storage of the basic OMB ethnicity categories (Hispanic or Latino, and Not Hispanic or Latino) does not use the unique 6-digit identifier codes. Rather, CIPHER approaches the collection of the minimum categories through use of a Yes/No data element. In contrast, collection/storage of the optional extended ethnicity categories associated with each of the minimum OMB categories involves the use of the unique 6-digit identifier codes described in the CDC/Census Ethnicity Classification table.

However, in the area of Electronic Data Interchange, both basic Ethnicity data and extended Ethnicity data are represented using the unique 6-digit identifier codes, to structure electronic messages that comply with HL7 and other communication standards.

EDI Summary

NOTE: All Ethnicity-related data will be formatted using the appropriate set of unique identifier codes noted in the CDC/Census Ethnicity Classification table prior to any type of electronic data interchange.

The remainder of the EDI section is under construction.

Implementation: Ethnicity

The implementation examples noted below specifically apply to the Ethnicity of the subject of the report. The implementation for other uses of Ethnicity, such as Ethnicity of the subject's partner, can be patterned after these implementation examples.

Data Collection: Hardcopy Report Form

Check-box fields on the hardcopy report form are used for the collection of Ethnicity data. The reporter can check or mark the box noting 'Y' for Hispanic or Latino or 'N' for Not Hispanic or Latino. Refer to Figures 1 and 2 below.

In addition, programs have the option to collect detailed extended Ethnicity data. The extended Ethnicity categories are described in the CDC/Census Ethnicity Classification table included at the end of this section. NOTE: A unique 6-digit identifier code is associated with each extended ethnicity category. Because the list is so extensive, an index of these standard extended ethnicity codes/categories will be accessible through a separate hardcopy reference. The reporter may choose to note the extended ethnicity category description and/or the extended ethnicity category unique identifier on the physical hardcopy form. Refer to Figures 3 through 5 for examples of hardcopy forms used to collect the optional extended Ethnicity data.

Figure 1:   **Blank Hardcopy Form section used to collect minimum Ethnicity data**

Ethnicity
(select one)

☐   **Y - Hispanic or Latino**

☐   **N - Not Hispanic or Latino**

Figure 2:   **Completed Hardcopy Form section used to collect Ethnicity data**

Ethnicity
(select one)

☑   **Y - Hispanic or Latino**

☐   **N - Not Hispanic or Latino**

Figure 3: **Blank Hardcopy Form section used to collect minimum Ethnicity data, and extended Ethnicity data**

Ethnicity
(select one)
☐   Y - Hispanic or Latino

Specify: _____

☐   N - Not Hispanic or Latino

Figure 4: **Completed Hardcopy Form section used to collect minimum Ethnicity data, and extended Ethnicity data using textual description**

Ethnicity
(select one)
☑   Y - Hispanic or Latino

Specify: Puerto Rican

☐   N - Not Hispanic or Latino

Figure 5: **Completed Hardcopy Form section used to collect minimum Ethnicity data plus extended Ethnicity data using code and description**

Ethnicity
(select one)
☑   Y - Hispanic or Latino

Specify: Puerto Rican   2180-8

☐   N - Not Hispanic or Latino
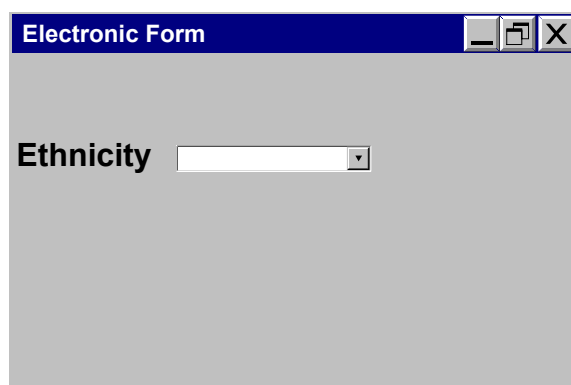
Missing Values – Hardcopy Form

Examples of hardcopy forms using the associated Missing Value Reason data element can be found in Appendix I – Missing Value Reason. The hardcopy form need only contain a missing value reason if the program requires the rationale for a missing value for Ethnicity.

Data Entry: Electronic Forms

A pull-down menu displays the valid entry options, which parallel the options noted on the collection report form ('Y' for Hispanic or Latino, 'N' for Not Hispanic or Latino). The portion of the entry options noted outside the parentheses reflects the data that are stored. Only the 'Y' or 'N' is stored in a database with the associated Ethnicity data element. Refer to Figures 6 through 8 below for examples of electronic entry forms used to enter the minimum Ethnicity concept data.
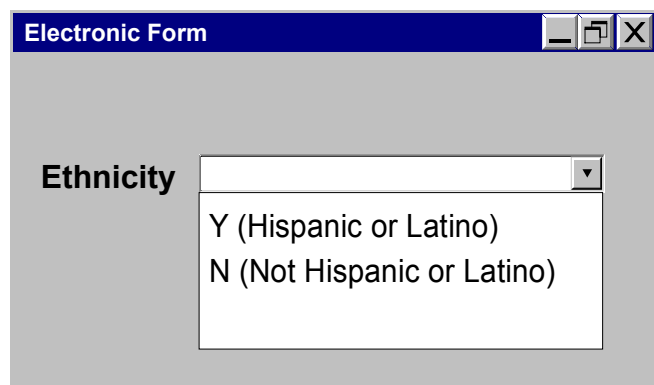
In addition, programs have the option to collect detailed extended Ethnicity information. The extended Ethnicity categories are described in the CDC/Census Ethnicity Classification table included at the end of this section. NOTE: A unique 6-digit identifier code is associated with each extended Ethnicity category. When a 'Y' is entered in the main Ethnicity field, indicating Hispanic or Latino, the extended Ethnicity data field is displayed along with a pull-down menu noting the valid entry options. These options parallel the options noted on in the CDC/Census Ethnicity Classification table. The portion of the valid entry options displayed outside the parentheses reflects data that are stored in the extended ethnicity data element, specifically the **6-digit unique identifier**. Refer to Figures 9 through 11 for examples of electronic forms used to collect the optional extended Ethnicity data.

Figure 6:   **Blank Electronic Form used to collect Ethnicity**
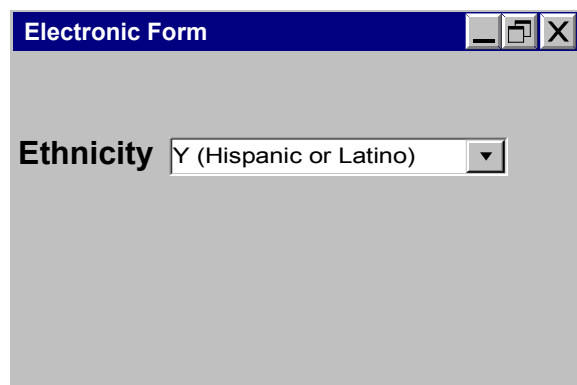
Figure 7:   **Blank Electronic Form used to collect Ethnicity, pull-down menu enabled**
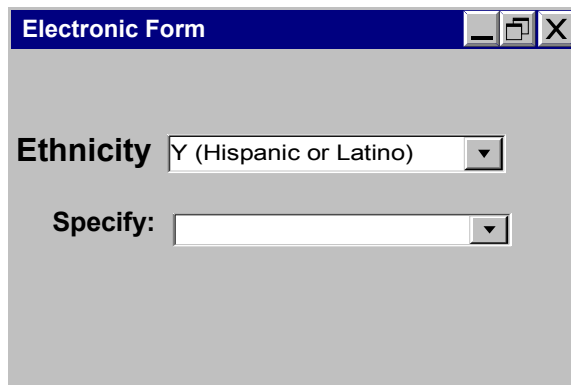
Electronic Form   _ 🗗 X

Ethnicity   [▾]

  Y (Hispanic or Latino)
  N (Not Hispanic or Latino)

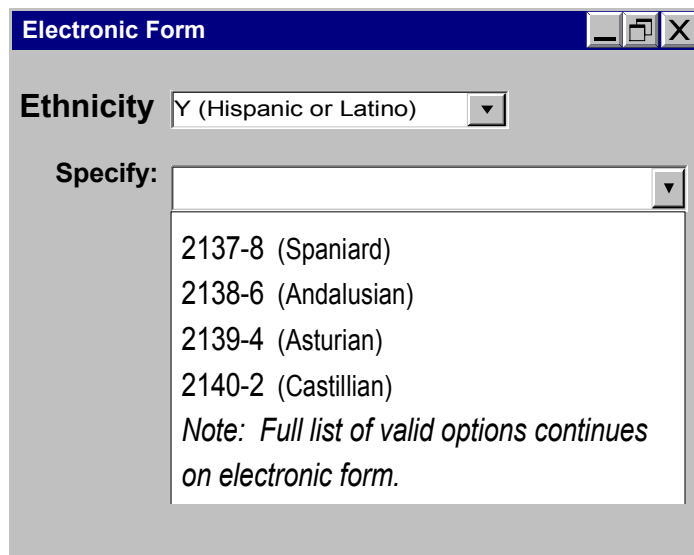**Figure 8:**   **Completed Electronic Form used to collect Ethnicity**

Electronic Form   _ 🗗 X

Ethnicity [Y (Hispanic or Latino)   ▾]

Figure 9:   **Electronic Form used to collect extended Ethnicity data**
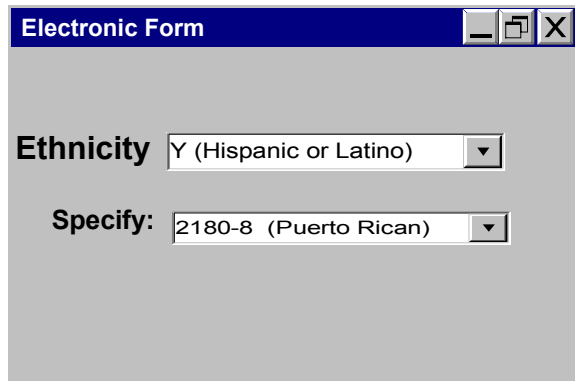
Electronic Form   ▢

Ethnicity Y (Hispanic or Latino) ▾

Specify: [   ] ▾

Figure 10:   **Blank Electronic Form used to collect extended Ethnicity data, pull-down menu enabled**

Electronic Form   ▢

Ethnicity Y (Hispanic or Latino) ▾

Specify: [   ] ▾

2137-8 (Spaniard)

2138-6 (Andalusian)

2139-4 (Asturian)

2140-2 (Castillian)

Note: Full list of valid options continues on electronic form.

Figure 11:   Completed Electronic Form used to collect extended Ethnicity data

Missing Values – Electronic Form

Examples of electronic forms using the associated Missing Value Reason (MVR) data element can be found in Appendix I – Missing Value Reason. The electronic form needs to handle the Missing Value Reason only if the program requires the reason for a missing value for Ethnicity. If the user selects an MVR code during data entry, the Ethnicity field will be blank and the screen will display the MVR information next to the blank field.

Data Processing: Validations and Edit Checks

Data elements entered in the electronic form will be edited as outlined below. If the program elects to use an associated Missing Value Reason data element for Ethnicity, it will be edited as outlined in Appendix I – Missing Value Reason.

Ethnicity:
- Alpha only
- A single Yes/No type variable is used to support the Ethnicity data concept. A value of 'Y' (Yes) indicates that the Ethnicity category is *Hispanic or Latino*. A value of 'N' (No) indicates that the Ethnicity category is *Not Hispanic or Latino*.
- **Ethnicity** data are to be collected in conjunction with **Race** data. Further, ethnicity data are to be noted and collected **before** Race data. Please refer to the Race section for detailed information on Race data.
- The extended Ethnicity data element (HISP_EXT) is activated when the following conditions are met:
  - a)   A program chooses to collect more detailed information on Ethnicity, and
  - b)   The associated Ethnicity data element (HISPANIC) has been marked with a 'Y' (Yes) indicating *Hispanic or Latino*.

Data Processing: From Hardcopy to Storage – minimum Ethnicity data

The following example illustrates the flow of information from data collection on the hardcopy form, to data entry into the electronic form, to validations and storage in the database.

The process begins with the blank Hardcopy data collection form used to collect Ethnicity:

The Ethnicity information is captured on the form, creating a completed Hardcopy data collection form:

The process continues with a blank Electronic form/data entry screen used to capture Ethnicity:

The value from the hardcopy form is entered into the Electronic form/data entry screen with the use of a drop-down list of valid values, and then the edits and validations are performed on Ethnicity:

The completed Electronic form/data entry screen is redisplayed and Ethnicity is stored in the database:

Electronic Form

Ethnicity: Y (Hispanic or Latino)

Database Storage

Variable:	*hispanic*
Type:	character
Length:	1
Stored Value:	**Y**

Data Processing: From Hardcopy to Storage – minimum Ethnicity data and extended Ethnicity data

The following example illustrates the flow of information from data collection on the hardcopy form, to data entry into the electronic form, to validations and storage in the database.

The process begins with the blank Hardcopy data collection form used to collect Ethnicity and extended ethnicity data:

The Ethnicity information is captured on the form, creating a completed Hardcopy data collection form:

The process continues with a blank Electronic form/data entry screen used to capture Ethnicity:

The value from the hardcopy form is entered into the Electronic form/data entry screen with the use of a drop-down list of valid values, and then the edits and validations are performed on Ethnicity:

The completed <u>Electronic form/data entry screen</u> is redisplayed and Ethnicity is stored in the database, along with extended ethnicity data:

Data Transmission: Electronic Data Interchange

NOTE: All Ethnicity-related data will be formatted using the appropriate set of unique identifier codes noted in the CDC/Census Ethnicity Classification table, prior to any type of electronic data interchange.

The remainder of the EDI section is under construction.

CDC/Census Ethnicity Classification

Unique Identifier	Hierarchical Code	Concept
2133-7	E	Ethnicity
2135-2	E1	Hispanic or Latino
2137-8	E1.01	Spaniard
2138-6	E1.01.001	Andalusian
2139-4	E1.01.002	Asturian
2140-2	E1.01.003	Castillian
2141-0	E1.01.004	Catalonian
2142-8	E1.01.005	Belearic Islander
2143-6	E1.01.006	Gallego
2144-4	E1.01.007	Valencian
2145-1	E1.01.008	Canarian
2146-9	E1.01.009	Spanish Basque
2148-5	E1.02	Mexican
2149-3	E1.02.001	Mexican American
2150-1	E1.02.002	Mexicano
2151-9	E1.02.003	Chicano
2152-7	E1.02.004	La Raza
2153-5	E1.02.005	Mexican American Indian
2155-0	E1.03	Central American
2156-8	E1.03.001	Costa Rican
2157-6	E1.03.002	Guatemalan
2158-4	E1.03.003	Honduran
2159-2	E1.03.004	Nicaraguan
2160-0	E1.03.005	Panamanian
2161-8	E1.03.006	Salvadoran
2162-6	E1.03.007	Central American Indian
2163-4	E1.03.008	Canal zone
2165-9	E1.04	South American
2166-7	E1.04.001	Argentinean
2167-5	E1.04.002	Bolivian
2168-3	E1.04.003	Chilean
2169-1	E1.04.004	Colombian
2170-9	E1.04.005	Ecuadorian
2171-7	E1.04.006	Paraguayan
2172-5	E1.04.007	Peruvian
2173-3	E1.04.008	Uruguayan
2174-1	E1.04.009	Venezuelan
2175-8	E1.04.010	South American Indian
2176-6	E1.04.011	Criollo
2178-2	E1.05	Latin American
2180-8	E1.06	Puerto Rican

2182-4	E1.07	Cuban
2184-0	E1.08	Dominican
2186-5	E2	Not Hispanic or Latino